Task 2. Remake the plot (35%)

The rise of the A.I. Large Language Models

4168216

The data visualisation website *Information Is Beautiful* contains a visualisation the rise of A.I. Large Language Models (LLMs). It shows the dates at which the models (and associated bots) were published, the number of parameters they are trained on, and which company owns them.

You can see the visualisation on this website. In the assignment you see a simplified version of this plot, e.g. indicating fewer owners and without labels for each model.

Your task is to remake the plot below. The variables that you will need are:

- `name`: name of the Large Language Model
- `owner_simplified`: company that owns the model (simplified version, contains fewer categories than original variable `owner`)
- `parameters`: number of parameters the model was trained on (in billions)
- `date`: date of publication of the model

Template file and submission:

Add your code to the provided template file. Write reproducible and readable code and make sure that the plot is visible in your output file (.html/.pdf). Keep the data stored in the subfolder so your .Rmd file can reach it.

For more submission instructions, please see the general instructions of this Graded Assignment.

Load data:

```
# Load packages

library(tidyverse)
library(foreign)

# Import data:
dataLLMs <- read_csv2("aiBasedLLMs.csv")

summary(dataLLMs)
```

```
##      name              owner             parameters          date
##  Length:125         Length:125         Min.   :   0.34   Min.   :2018-10-01
##  Class :character   Class :character   1st Qu.:  11.00   1st Qu.:2022-04-01
##  Mode  :character   Mode  :character   Median :  52.00   Median :2023-03-01
##                                        Mean   : 132.89   Mean   :2022-11-23
##                                        3rd Qu.: 175.00   3rd Qu.:2023-09-01
##                                        Max.   :1750.00   Max.   :2024-03-01
##  owner_simplified
##  Length:125
##  Class :character
```

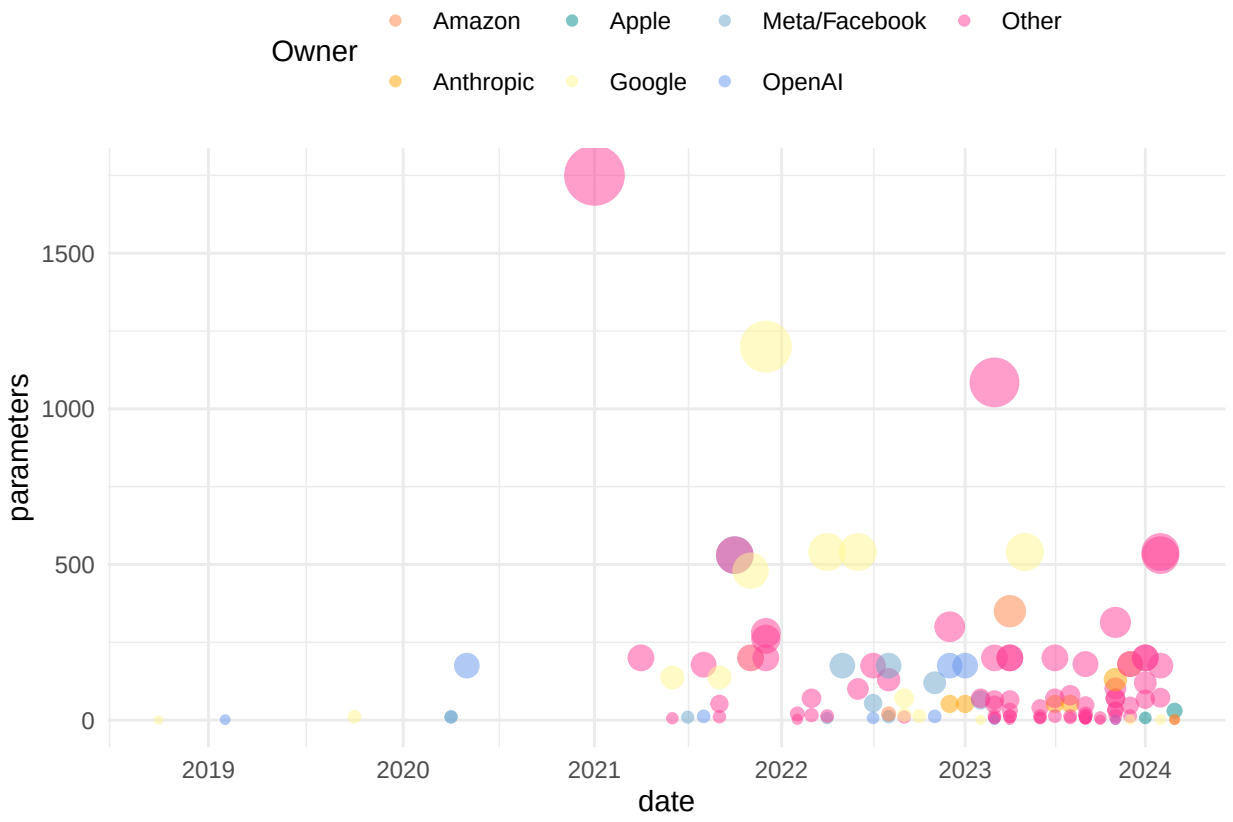
```
##  Mode  :character
##
##
##
```

```
#ggplot(dat)
```

```
ggplot(dataLLMs,
       aes(x = date,
           y = parameters,
           size = parameters,
           color = owner_simplified
           )) +
  geom_point(alpha = 0.5) +
  scale_radius(name = "Body Mass (g)") +
  scale_size_continuous(range = c(1, 10)) +
  theme(legend.position = "top") + theme_minimal() +  theme(legend.position="top") + guides(fill=guide_l
  scale_x_date(date_breaks="1 year", date_labels="%Y") +
  coord_trans(x="log2") + guides(size=FALSE) + labs(color = "Owner") +
   scale_color_manual(values = c("Amazon" = "sienna1",
                                 "Apple" = "darkcyan",
                                 "Anthropic" = "orange1",
                                 "Google" = "khaki1",
                                 "OpenAI" = "cornflowerblue",
                                 "Meta/Facebook" = "skyblue3",
                                 "Other" = "violetred1"  ))
```

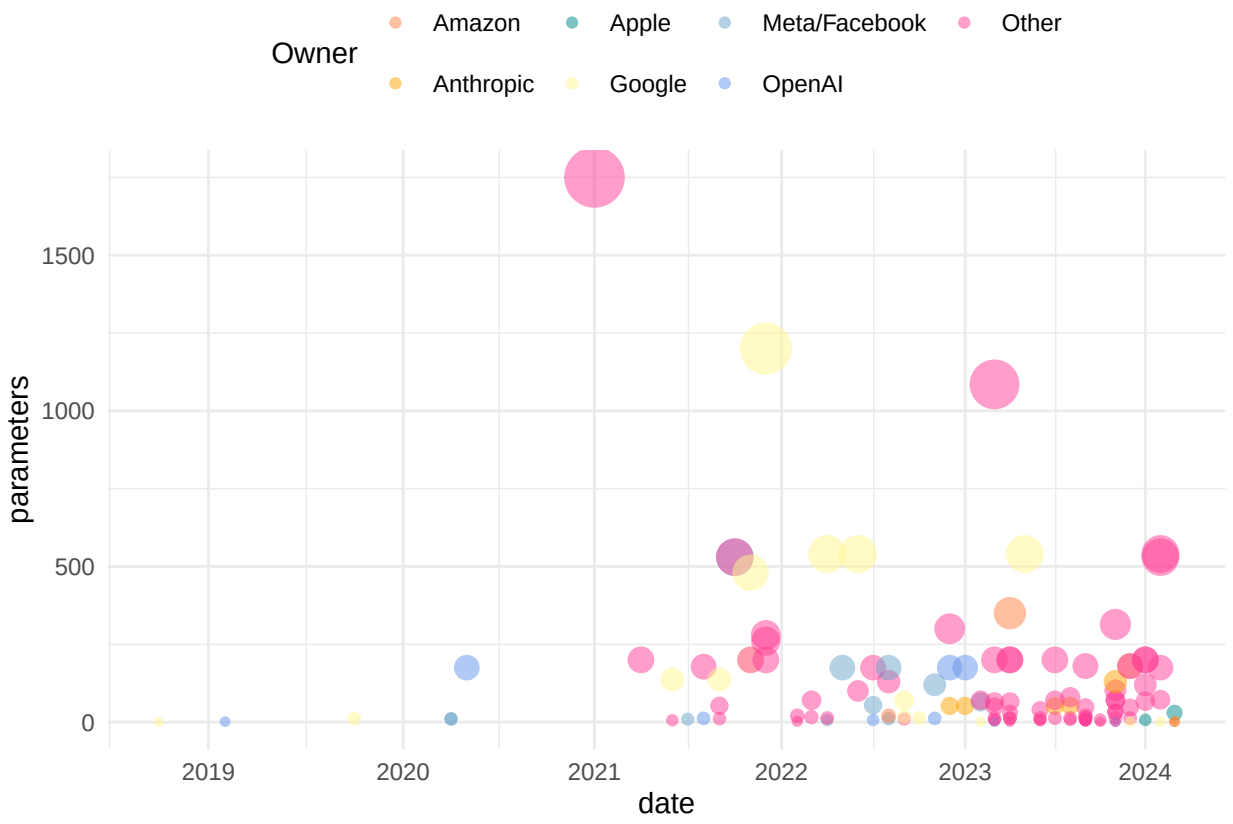
```
## Scale for size is already present.
## Adding another scale for size, which will replace the existing scale.
```

```
## Warning: The `<scale>` argument of `guides()` cannot be `FALSE`. Use "none" instead as
## of ggplot2 3.3.4.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```




```
ggplot(dataLLMs,
       aes(x = date,
           y = parameters,
           size = parameters,
           color = owner_simplified
           )) +
  geom_point(alpha = 0.5) +
  scale_radius(name = "Body Mass (g)") +
  scale_size_continuous(range = c(1, 10)) +
  theme(legend.position = "top") + theme_minimal() +  theme(legend.position="top") + guides(fill=guide_l
  scale_x_date(date_breaks="1 year", date_labels="%Y") +
  coord_trans(x="log2") + guides(size=FALSE) + labs(color = "Owner") +
   scale_color_manual(values = c("Amazon" = "sienna1",
                                 "Apple" = "darkcyan",
                                 "Anthropic" = "orange1",
                                 "Google" = "khaki1",
                                 "OpenAI" = "cornflowerblue",
                                 "Meta/Facebook" = "skyblue3",
                                 "Other" = "violetred1"  ))
```

```
## Scale for size is already present.
## Adding another scale for size, which will replace the existing scale.
```



Your plot:

```
# Do not change the above fig.width = 8 and fig.height = 6; it will give you the same dimensions as the

# your code here
```